Some Remarks on Approximate Entailment

E. Trillas
Facultad Informática,
Universidad Politécnica de Madrid,
Madrid, Spain

C. Alsina
E.T.S. Arquitectura Barcelona,
Universidad Politécnica de Catalunya,
Barcelona, Spain

ABSTRACT

Fuzzy logic states based upon a very general rule of inference are introduced and studied with detail. Relations with fuzzy preorders are given. In particular, several residuated implications used in expert systems are justified from a theoretical point of view, and some special cases related to probabilistic models are analyzed.

KEYWORDS: ***fuzzy logic state, fuzzy preorder, entailment, approximate reasoning, residuated implications, t-norms, implications, approximate consequence, conditional probability***

Motivated by some problems of inference arising in the field of expert systems and concerning the implementation of the modus ponens rule, we study in this paper some structures called fuzzy logic states, which are weaker than fuzzy preorders but closely related to them, and we exhibit their precise relations. In particular, we point out that some residuated implications that have been used on an empirical basis may be justified theoretically as being "better" objects than some classical probabilistic parameters, and we study in detail cases where some classes of residuated implications may be more convenient than others. We also introduce the concept of approximate consequences and their relations with our model of fuzzy logic states, and finally, we study some relations of our results in the context of nonparametric statistics.

Address correspondence to David Sher, Computer Science Department, SUNY Buffalo, Buffalo, NY 14260.
Received May 1, 1991; accepted November 8, 1991.

International Journal of Approximate Reasoning 1992; 6:525–533

655 Avenue of the Americas, New York, NY 10010 0888-613X/92/$5.00

FUZZY LOGIC STATES AND FUZZY PREORDERS

DEFINITION 1 *A* fuzzy logic state, *or briefly a* fuzzy state, *is given by* (F, R, μ, M), *where F is a given (nonempty) set,* $R: F \times F \to [0, 1]$ *is a fuzzy relation,* $\mu: F \to [0, 1]$ *is a fuzzy subset, and* $M: [0, 1] \times [0, 1] \to [0, 1]$ *satisfies* $M(1, 1) = 1$, *and, for all a, b in F,*

$$M(\mu(a), R(a, b)) \leq \mu(b). \tag{1}$$

Thus, in a fuzzy logic state, if $\mu(a) = 1$ and $R(a, b) = 1$, then $\mu(b) = 1$; that is, we have a rule of inference that is a weak version of the classical modus ponens. Moreover, if M is a *positive* two-place function [i.e., $M(x, y) > 0$ whenever $x > 0$ and $y > 0$], then by (1) we obtain that μ is a *fuzzy set of true elements* (Trillas and Alsina [1]) [i.e., if $\mu(a) > 0$ and $R(a,b) > 0$, we deduce that $\mu(b) \geq M(\mu(a), R(a, b)) > 0$].

A natural way to define fuzzy logic states is to use fuzzy preorders.

DEFINITION 2 *A* fuzzy preorder *in a set F is a mapping* $I: F \times F \to [0, 1]$ *such that* $I(a \mid a) = 1$ *for all a in F (reflexivity) and there exists a binary operation M on* $[0, 1]$ *such that* $M(1, 1) = 1$ *and for any a, b, c in F we have the transitivity property*

$$M(I(c \mid a), I(a \mid b)) \leq I(c \mid b). \tag{2}$$

It is immediate from the above definitions that if (F, I, M) is a fuzzy preorder, then for any c in F if we consider $\mu_c: F \to [0, 1]$ given by $\mu_c(x) = I(c \mid x)$, then (F, I, μ_c, M) is a fuzzy state with $R(a, b) = I(a \mid b)$. In the other direction, let us begin with a fuzzy state (F, R, μ, M), where $M(x, y) = m^{(-1)}[m(x) + m(y)]$ is an Archimedean t-norm [with m: $[0, 1] \to [0, +\infty]$ continuous, strictly decreasing, $m(1) = 0$, and $m^{(-1)}(x) = m^{-1}(x)$ if $x \leq m(0)$ and $m^{(-1)}(x) = 0$ if $x > m(0)$]. Then (Trillas and Alsina [1]) we have

$$R(a, b) \leq m^{(-1)} \operatorname{Max}\{m(\mu(b)) - m(\mu(a)), 0\} := I_M(b \mid a), \tag{3}$$

and if we consider the right-hand side of (3) and define

$$I_M(b \mid a) := m^{(-1)} \operatorname{Max}\{m(\mu(b)) - m(\mu(a)), 0\},$$

then (F, I_M, M) is a fuzzy preorder relative to M. Thus our initial fuzzy relation becomes bounded from above by the fuzzy preorder (Trillas and Valverde [2]) associated to the t-norm M [in fact, $I_M(b \mid a) = \sup\{z \in [0, 1] \mid M(\mu(b), z) \leq \mu(a)\}$].

Note that when there exists some prototype $b_0 \in F$, that is, $\mu(b_0) = 1$, then $I_M(b_0 \mid a) = \mu(a)$. This is the case in both conditional probability and probability of a material implication. When $M = \operatorname{Prod}[(x, y) = x \cdot y]$ or $M =$

W [the Lukasiewicz t-norm $W(x, y) = \mathrm{Max}(x + y - 1, 0)$], we have, respectively,

$$I_{\mathrm{Prod}}(b \mid a) = \begin{cases} \mathrm{Min}[1, \mu(b)/\mu(a)] & \text{if } \mu(a) \neq 0, \\ 1 & \text{if } \mu(a) = 0, \end{cases} \tag{4}$$

and

$$I_W(b \mid a) = \mathrm{Min}[1, 1 - \mu(a) + \mu(b)]. \tag{5}$$

Both fuzzy preorders (4) and (5) are well-known *residuated implication operators* (see Trillas and Valverde [37]) widely used in expert systems: (4) is the *Gaines 43* operator and (4) is the *Lukasiewicz* operator.

NOTE If $(F, R, \mu, \mathrm{Min})$ is a fuzzy state, then instead of (3) we have

$$R(a, b) \leq R_1(a, b) := \begin{cases} 1 & \text{if } \mu(a) \leq \mu(b), \\ \mu(b) & \text{if } \mu(a) > \mu(b), \end{cases}$$

and R_1 is a fuzzy preorder with respect to Min, the *standard star* operator.

In effect, if $\mu(a) \leq \mu(b)$, then it is obvious that $R(a, b) \leq 1 = R_1(a, b)$. When $\mu(a) > \mu(b)$, the possibility $R(a, b) > \mu(b)$ would yield

$$\mu(b) = \mathrm{Min}[\mu(a), \mu(b)] \leq \mathrm{Min}[\mu(a), R(a, b)] \leq \mu(b),$$

and therefore we would obtain the chain of equalities

$$\mathrm{Min}[\mu(a), \mu(b)] = \mathrm{Min}[\mu(a), R(b \mid a)] = \mu(b);$$

that is, either $\mu(a) = \mu(b)$ or $R(b \mid a) = \mu(b)$, which is a contradiction. Thus $R(b \mid a) \leq \mu(b)$ whenever $\mu(a) > \mu(b)$. Moreover, the proof that R_1 is a fuzzy preorder with respect to Min is a straightforward computation. Preorder R_1, the Gödel operator, is the greatest residuated implication (Trillas and Valverde [3]).

One of the reasons I_W appears so frequently in the literature concerning theoretical aspects of expert systems is contained in the following result. Consider a set F of facts where there is a structure $(F, +, \cdot, ', \leq)$ corresponding to OR, AND, NOT, and IMPLIES, respectively, and v: $F \to [0, 1]$ is supposed to satisfy the following conditions.

- **(i)** If $a \leq b$ then $v(a) \leq v(b)$.
- **(ii)** $v(a') = 1 - v(a)$.
- **(iii)** $v(a \cdot b) = T(v(a), v(b))$ and $v(a + b) = 1 - T(1 - v(a), 1 - v(b))$ for some t-norm T.
- **(iv)** $v(a + b) + v(a \cdot b) = v(a) + v(b)$.

If in (F, v) we assume these last conditions and we define $R_v^T\colon F \times F \to [0, 1]$ by

$$R_v^T(a, b) := v(a' + b) = v(a') + v(b) - v(a' \cdot b)$$
$$= 1 - v(a) + v(b) - T(1 - v(a), v(b)),$$

then we have, in the case that $a \cdot (a' + b) \leq a \cdot b \leq b$,

$$T(v(a), R_v^T(a, b)) = T(v(a), v(a' + b)) = v(a \cdot (a' + b))$$
$$\leq v(a \cdot b) \leq v(b),$$

that is, (F, R_v^T, v, T) is a fuzzy state. Thus, it would be interesting to know when R_v^T is in fact a fuzzy preorder. The reflexivity $R_v^T(a, a) = 1$ for all $a \in E$ yields that necessarily

$$1 = R_v^T(a, a) = 1 - T(1 - v(a), v(a)),$$

that is,

$$T(1 - v(a), v(a)) = 0,$$

and consequently if there is a $v(a) \in (0, 1)$ we have that T cannot be either Min or a strict t-norm. In the case that $v\colon F \to [0, 1]$ is onto and T is continuous, we deduce from (iii) and (iv) that the t-norm T must satisfy the well-known Frank's equation

$$T(x, y) + 1 - T(1 - x, 1 - y) = x + y,$$

and the unique continuous solution satisfying the requirement $T(x, 1 - x) = 0$ is precisely $T = W$. But in this case we have

$$R_v^W(a, b) = 1 - v(a) + v(b) - W(1 - v(a), v(b))$$
$$= 1 - v(a) + v(b) - \mathrm{Max}\left[v(b) - v(a), 0\right]$$
$$= \mathrm{Min}\left[1, 1 - v(a) + v(b)\right] = I_W(a \mid b).$$

Thus I_W arises naturally in this context, and no other fuzzy preorders would have sense.

NOTE 1 R_v^{Min} and R_v^{Prod} satisfy neither the reflexivity condition nor the transitivity inequality, and even the weak modus ponens "$R(a, b) > 0$ and $v(a) > 0$ implies $v(b) > 0$" is not satisfied.

NOTE 2 In order to derive (3) from (1) it was a key point that for an Archimedean t-norm M it was possible to solve the inequality $M(x, z) \leq y$

for fixed x and y. A similar situation arises when M is a quasi-arithmetic mean, that is, (3) is still possible in other algebraic structures that are nonassociative (e.g., M is the geometric mean).

NOTE 3 In many cases, it is possible to have a fuzzy logic state and not a fuzzy preorder. Actually, the inequality (1) is used to get an idea about possible values of $\mu(b)$ by knowing the value of $\mu(a)$ and the values $R(a, b)$ expressing the strength of the entailment between a and b. That is, if μ is known for some $K \subset F$ and we have R on $F \times F$, then by (1) we can be sure that even for $b \in F \setminus K$ we will have $\mu(b) \in [M(\mu(a), R(a, b)), 1]$. Thus by virtue of (1) we may define μ on $F \setminus K$ by $\mu(b) := \sup\{M(\mu(x), R(b, x)) \mid x \in F, R(b \mid x) > 0\}$ and so extend the partial function μ to new values. In many measure considerations one may have $\mu(a, b) = M(\mu(a), R(a, b))$, and by the monotonicity of the measure we deduce (1).

NOTE 4 It is well known (López de Mántaras [4]) that if N is a necessity measure and the law $a \cdot b = a \cdot (a' + b)$ holds in F, then $N(b) \geq \mathrm{Min}[N(a), N(a \to b)]$ and therefore $(F, N(\to), N, \mathrm{Min})$ is a fuzzy logic state.

FUZZY STATES AND APPROXIMATE CONSEQUENCES

In [1] it was shown that, given a fuzzy preorder I on F, the operator in the power set of F, $\mathscr{P}(F)$, defined via

$$C(A) = \{b \in F;\ I(a \mid b) > 0 \text{ for some } a \in A\},$$

is a consequence operator in the classical sense of Tarski. Conversely, it is easy to derive classical preorders from consequence operators.

Let us assume that for ϵ in $(0, 1)$ we define, from a fuzzy preorder (F, I, M), the operator C^ϵ in $\mathscr{P}(F)$ by

$$C^\epsilon(A) = \{b \in F;\ I(a \mid b) > \epsilon \text{ for some } a \in A\}.$$

If M is a strict t-norm, then $C^\epsilon(\{a\})$ coincides with the open ball of center a and radius $1 - \epsilon$ in the generalized metric space (Schweizer and Sklar [5]) given on F by the nonreflexive generalized metric $1 - I(a \mid b)$, and elements of $C^\epsilon(\{a\})$ could be called *ϵ-approximate consequences* of a (their "logical distances" to a are less than or equal to $1 - \epsilon > 0$).

It is clear that $A \subset C^\epsilon(A)$ because $I(a \mid a) = 1 > \epsilon$, and if $A \subset B$, then $C^\epsilon(A) \subset C^\epsilon(B)$. If M is nondecreasing in both arguments, we also obtain $C^\epsilon(C^\delta(A)) \subset C^{M(\epsilon, \delta)}(A)$. Needless to say, if M is positive, $M(\epsilon, \delta) > 0$, but even for nonpositive M (e.g., $M = W$) we still may find values ϵ, δ

for which $M(\epsilon, \delta) > 0$. Thus. when $M(\epsilon, \delta) > 0$ we obtain $C^{M(\epsilon, \delta)}(A) \subset C(A)$. For example, if $M = W$ and $\epsilon > 1/2$, we have $C^{\epsilon}(C^{\epsilon}(A)) \subset C^{W(\epsilon, \epsilon)}(A) \subset C(A)$. The importance of the inclusions $C^{\epsilon}(C^{\delta}(A)) \subset C^{M(\epsilon, \delta)}(A) \subset C(A)$ is that the ϵ-approximate consequences of the δ-approximate consequences of A are consequences of A.

Thus, even if we start with a fuzzy state and construct, the fuzzy preorder I_M, then, by applying the machinery just described, we can have a logical background in terms of either Tarski's operators or ϵ-approximate consequences.

FUZZY STATES AND PROBABILITIES

Letting $(F, +, \cdot, ')$ be a Boolean algebra, p: $F \to [0, 1]$ is a probability measure on F, and $F^+ = \{a \in F \mid p(a) > 0\}$. As was seen in [1], the *conditional probability* P: $F \times F^+ \to$ [0, 1] defined by

$$P(b \mid a) := \frac{p(a \cdot b)}{p(a)} \tag{6}$$

does not determine a fuzzy preorder on F^+ of any operation M with 1 as a unit. Moreover, the *probability of a material implication*,

$$p(a \to b) := p(a' + b), \tag{7}$$

does not induce a fuzzy preorder on F for M = Prod or M = Min (but it defines a fuzzy preorder for the nonpositive t-norm $M = W$).

Thus, in the probabilistic Boolean realm, the usual, and useful, relations (6) and (7) do not yield fuzzy preorders under positive operations. Nevertheless, we can capture such structures under the framework of fuzzy states.

PROPOSITION 1 *The conditional probability* (6) *satisfies* (1) *with* M = Prod, *and therefore* (F^+, P, p, Prod) *is a fuzzy state. The probability of a conditional* (7) *satisfies* (1) *under* $M = W$, *and* (F, $p(\to)$, p, W) *is a fuzzy state.*

Proof The first part of the statement follows from the monotonicity of the measure p and the inequality $p(a) \cdot P(b \mid a) = p(ab) \leq p(b)$. For the other case, we need to take into account the fact that

$$\begin{aligned} W(p(a), p(a \to b)) &= W(p(a), 1 - p(a) + p(b) - p(a' \cdot b)) \\ &= \mathrm{Max}[0, p(b) - p(a' \cdot b)] \leq p(b). \end{aligned}$$

In view of the above proposition and (3), $P(b \mid a)$ and $p(a \to b)$ will have as upper bounds the fuzzy preorders I_{Prod} and I_W; that is,

$$P(b \mid a) \leq I_{\mathrm{Prod}}(a \mid b) \qquad \text{and} \qquad p(a \to b) \leq I_W(a \mid b),$$

with I_{Prod} and I_W as given by (4) and (5), respectively. It is interesting to compute the differences

$$I_{\text{Prod}}(a \mid b) - P(b \mid a) = \frac{p(a+b) - \text{Max}[p(a), p(b)]}{p(a)}$$
$$= \frac{\text{Min}[p(a), p(b)] - p(a \cdot b)}{p(a)}, \tag{8}$$

and

$$I_W(a \mid b) - p(a \to b) = \begin{cases} p(a \cdot b') & \text{if } p(a) \leq p(b), \\ p(a' \cdot b) & \text{if } p(a) \geq p(b). \end{cases} \tag{9}$$

Equations (8) and (9) express how far the relations $P(b \mid a)$ and $p(a \to b)$ are from the fuzzy preorder values $I_{\text{Prod}}(a \mid b)$ and $I_W(a \mid b)$, respectively. Note that $P(b \mid a) = I_{\text{Prod}}(b \mid a)$ if and only if $p(a \cdot b) = \text{Min}[p(a), p(b)]$, a condition used in some expert systems (López de Mántaras [4]) that is clearly satisfied whenever a and b are comparable in the Boolean order, that is, $a \leq b$ $(a \cdot b = a)$ or $b \leq a$ $(a \cdot b = b)$, and it can be seen immediately that the difference (8) becomes small whenever $p(a)$ is close to 1, that is,

$$0 \leq I_{\text{Prod}}(a \mid b) - P(b \mid a) \leq \frac{\text{Min}[p(a), p(b)] - W(p(a), p(b))}{p(a)}$$
$$\leq \frac{\text{Min}\{\text{Min}[p(a), p(b)], 1 - \text{Max}[p(a), p(b)]\}}{p(a)}$$
$$\leq \frac{1 - p(a)}{p(a)}.$$

Clearly, if $I_W(a \mid b) = p(a \to b)$, then $p(a \cdot b') \cdot p(a' \cdot b) = 0$.

It is evident from our consideration that in most common cases $I_{\text{Prod}}(a \mid b)$ will be a "better" relation that $P(b \mid a)$ because it is the fuzzy preorder generated by the probability p, a Tarski's consequence structure is obtained, and I_{Prod} is functionally expressable in terms of $p(a)$ and $p(b)$, whereas $P(b \mid a)$ does not enjoy such properties in general. A similar situation arises in the case I_W.

The difference (8) may be of some interest in nonparametric statistics. Precisely, let x, y random variables defined on a common probability space $(\Omega, \mathscr{F}, p)$ with values in $\mathbb{R}^+$ and with joint distribution function H_{xy} and marginals F_x and F_y. For any u, v in $\mathbb{R}^+$, consider the elements of $\mathscr{F}$:

$$a_x(u) = \{w \in \Omega \mid X(w) < u\} \qquad \text{and} \qquad a_y(v) = \{w \in \Omega \mid Y(w) < u\}.$$

Then, $p(a_x(u)) = F_x(u)$, $p(a_y(v)) = F_y(v)$, and (Schweizer and Sklar [5]) there exists a two-dimensional copula C_{xy} such that

$$H_{xy}(u, v) = p(a_x(u) \cap a_y(v)) = C_{xy}(F_x(u), F_y(v)).$$

Then we have, for the function

$$D_{xy}(u, v) = I_{\text{Prod}}(a_y(v) a_x(u)) - P(a_y(v) \mid a_x(u)),$$

that, in view of (8), if $F_x(u) \neq 0$ for all $u > 0$,

$$D_{xy}(u, v) = \frac{\operatorname{Min}[F_x(u), F_y(v)] - C_{xy}(F_x(u), F_y(v))}{F_x(u)}. \tag{10}$$

Thus, $D_{xy} \equiv 0$ if and only if x and y are positively linearly dependent ($y = Kx$, $K > 0$), that is, $(x, y) = +1$, or positively correlated. In the case where F_x and F_y are strictly increasing and continuous on $(0, +\infty)$ with $F_x(0) = F_y(0) = 0$ and $F_x(+\infty) = F_y(+\infty) = 1$, by introducing in (10) the change of variables $s = F_x(u)$, $t = F_y(v)$ and considering the function $L_{xy}(s, t) = D_{xy}(F_x^{-1}(s), F_y^{-1}(t))$, we obtain

$$L_{xy}(t) = \frac{\operatorname{Min}(s, t) - C_{xy}(s, t)}{s}.$$

Thus,

$$C_{xy}(s, t) = \operatorname{Min}(s, t) - sL_{xy}(s, t),$$

and L_{xy} shares with the copula C_{xy} intrinsic information concerning the dependence of x and y; for example, L_{xy} will be invariant under monotonic transformations of the random variables.

In relation with nonparametric measures of dependence, we obtain, for example, that the Spearman measure s as well as the Schweizer–Wolff measure (Schweizer and Sklar [5]) can be represented in terms of L_{xy}:

$$\rho = 12 \iint_{[0,1]^2} C_{xy}(s, t)\, ds\, dt - 3 = 1 - \iint_{[0,1]^2} sL_x(s, t)\, ds\, dt,$$

$$\sigma = 12 \iint_{[0,1]^2} |C_{xy}(s, t) - st|\, ds\, dt$$

$$= 12 \iint_{[0,1]^2} |\operatorname{Min}(s, t) - st - sL_{xy}(s, t)|\, ds\, dt,$$

whence X and Y are independent if and only if $\sigma = 0$, and this happens if and only if

$$L_{xy}(s, t) = \frac{\mathrm{Min}(s, t) - st}{s} = \mathrm{Min}\left(1, \frac{t}{s}\right) - t.$$

A more detailed study of L_{xy} in the context of nonparametric statistics is merited.

ACKNOWLEDGMENT

We thank the referees for their remarks, which made possible some improvements in the original manuscript.

References

1. Trillas, E., and Alsina, C., Logic: going further from Tarski, *Fuzzy Sets Syst.*, 1992.
2. Trillas, E., and Valverde, L., An inquiry into indistinguishability operators, *Aspects of Vagueness*, H. Skala et al., eds., Reidel, Dordrecht, 1984, pp. 231–256.
3. Trillas, E., and Valverde, L., On implication and indistinguishability in the settings of fuzzy logic, *Management Decision Support Systems Using Fuzzy Sets and Possibility Theory*, J. Kayvrvjgk and R. R. Yager, eds., Verlag TUV Rheinland, 1985.
4. López de Mántaras, R., *Approximate Reasoning Models*, Ellis Horwood, Chichester, 1990.
5. Schweizer, B., and Sklar, A., *Probabilistic Metric Spaces*, Elsevier North-Holland, New York, 1983.